

- 1 -

Sequence Listings

<110> Roche Vitamins AG
 <120> A gene encoding aldehyde dehydrogenase and use thereof
 <130> 21424
 5 <140>
 <141>
 <160> 9
 <170> PatentIn Ver. 3.1

10 <210> 1
 <211> 3408
 <212> DNA
 <213> Gluconobacter oxydans
 <220>

15 <221> CDS
 <222> (258) .. (2087)
 <220>
 <221> CDS
 <222> (2214) .. (3317)

20 <400> 1
 GCGACTGGCA GCAGCGCAAC TATGACCACT ATGGCCTGCC GCCCTATTGG 50
 ATCTAACTGA TCCAGTAAGC CACCATCAGC CGGCCCCCTGC GGGGGCCGGC 100
 TTTTGTGCGT AGACCCCGCC GAGGTGCTGT CGTAACCTAA GGTCACATCT 150
 25 TTACTTCCAC ATCCGCCCTT GTCAGTTCTG ACGTGACAAA TTGTCGCGGT 200
 CATGCTGCTG AATGCGGATG CCAGTCCCAG ATCCAAGCCC GACGCAAGGA 250
 GACGTAGATG TTACCCAAAT CATTGAAACA TAAGAATGGC GCCATGCGCC 300
 TTGTCGCAGC CTCGACCCTT GCGCTGATGA TCGGCGCGGG TGCCCATGCG 350
 CAGGTAAACC CGGTCGAAGT GCCGGTGGGC GCGAACGAGA CCTTTACCTC 400
 30 GCGCGTGCTG ACCACCGGCC TGTCTGAACCC TTGGGAAATC ACCTGGGGCC 450
 CCGACAATAT GCTGTGGGTG ACCGAGCGAT CTTCGGGCGA AGTGACGCGC 500
 GTCGACCCCA ATACCGGCGA GCAGCAGGTC CTGCTGACCC TGACCGATTT 550
 CAGCGTCGAT GTGCAACACC AGGGCCTACT TGGCCTCGCG CTGCATCCTG 600
 AGTTTATGCA AGAGAGCGGC AACGACTACG TCTATATCGT CTACACTTAT 650
 35 AACACCGGCA CCGAAGAAGC GCCCGATCCG CATCAAAAGC TGGTGCGTTA 700
 TGCCTATGAC GCTGCCGCGC AGCAGCTGGT CGATCCGGTT GATCTGGTCG 750
 CAGGCATTCC CGCAGGCAAC GACCACAATG GCGGTCGCAT CAAATTCGCC 800
 CCCGATGGCC AACACATCTT TTACACGCTG GGCGAGCAAG GCGCGAACTT 850
 TGGCGGTAAC TTCCGCCGTC CGAACCACGC GCAACTGCTG CCGACGCAAG 900
 40 AGCAGGTCGA CGCGGGCGAT TGGGTCGCCT ATTCGGGCAA GATCCTGCGC 950

GTGAACCTTG ACGGCACGAT CCCCGAAGAC AACCCCGAGA TCGAGGGCGT 1000
 GCGTAGCCAT ATCTTTACCT ATGGCCACCG TAACCCGCAG GGCATCACCT 1050
 TTGGCCCCGA CGGCACCATT TATGCCACCG AACACGGCCC CGATACGGAT 1100
 GACGAGCTGA ACATCATCGC CGGCGGTGGC AACTATGGGT GGCCGAATGT 1150
 5 GGCCGGCTAT CGCGATGGCA AATCTATGT CTACGCTGAT TGGAGCCAAG 1200
 CGCCCGCTGA CCAGCGTTAC ACCGGTCGCG CCGGTATCCC CGACACCGTG 1250
 CCGCAATTCC CCGAGCTGGA ATTCGCGCCC GAGATGGTCG ATCCGCTGAC 1300
 AACCTATTGG ACGGTGGATA ATGATTACGA TTTCACCGCC AATTGCGGCT 1350
 GGATCTGTAA TCCGACGATC GCGCCTTCGT CTGCCTATTA CTATGCGGCG 1400
 10 GGCGAGAGCG GTATCGCGGC TTGGGATAAT TCGATCCTGA TCCCGACGCT 1450
 GAAACATGGC GGCATCTATG TGCAGCACCT CAGCGATGAT GGCCAATCTG 1500
 TCGACGGCCT GCCCCAGCTG TGGTTCAGCA CCCAGAACCG CTATCGCGAT 1550
 ATCGAGATCA GCCCCGATAA CCATGTTTTT GTGGCGACCG ACAACTTTGG 1600
 CACCTCGGCG CAGAAATATG GCGAGACCGG CTTTACCAAC GTGCTGCATA 1650
 15 ACCCCGGCGC GATCCTTGTC TTTAGCTATG TCGGCGAGGA TGCTGCGGGT 1700
 CAGACCGGAA TGATGACCGC GCCCGCACCG CAGACGCAAT ACACGCAAGT 1750
 GCCCCCGAG GGTGCAGGCG CGGGCGCGAC TGAGGTTGCG GATGTCGATT 1800
 ACGACACGCT GTTCACCGAA GGCCAGACCC TTTATGGCAG CGCATGTGCC 1850
 GCGTGCCATG GTGCCGCTGG CCAAGGTGCG CAGGGCCCGA CCTTTGTGGG 1900
 20 CGTGCCGGAT GTGACGGGTG ACAAGGACTA CCTTGCCCGC ACCATCATCC 1950
 ACGGTTTTGG CTATATGCCG TCGTTGCGA CTCGGCTGGA TGACGAGGAG 2000
 GTTGCCGCCA TCGCGACCTT TATCCGAAC AGCTGGGGCA ATGACGAAGG 2050
 CATCCTGACC CCGGCCGAGG CCGCTGCCAC CCGCTGAATG CTGTAAAAAC 2100
 CACCCTCGCC TGCACATCAG GCGGGGGTAT TTCATTTATT TTCACATCTG 2150
 25 CCTTTGACAT GTGCCGCTAT CACGGTTAAT GCGGCCCTTC GGCTGTTCTG 2200
 GGTCTAAGCG GGTGTGTTGC CCGATAAGAG AGACGGTTCA GTCCCTCCCG 2250
 CCCTATTTAG GGCCCATTTA GGCGAGAATAG TTTTGACTCA TCAAAATATC 2300
 GCCGCGCCTC TGGCCGCGGC CCTTTCGCAA CGTGGATATG AAACGCTGAC 2350
 CGCCGTGCAG CAAGCTGTGC TTGCGCCCGA GGCTGATGGC CGCGACCTGC 2400
 30 TGGTGTGCGC ACAGACCGGT TCGGGTAAGA CGGTGGCCTT TGGTATCGCA 2450
 GTCGCGCCCG ACCTTTTGGG CGACGACAAT ATCCTGCCGC TGAACACGCC 2500
 GCCTGTTGCG CTGTTTCATCG CCCCCACGCG CGAGCTTGCG CTGCAAGTTG 2550
 CTCAGGAAC TACCTGGCTT TACGCCAATG CAGGTGCCCA GATCGCGACC 2600
 TGCGTCGGCG GTATGGATTA CCGCACCGAG CGCCGCGCCC TTGCACGTCT 2650
 35 GCCGCAAATC GTTGTGCGCA CGCCCGGCCG TCTGCGCGAC CATATCGACC 2700
 GTGGCGGCCT TGACCTGTCC GAATTGCGCG TGACCGTGCT GGACGAAGCG 2750
 GATGAGATGC TCGACCTCGG CTTCCGCGAT GATCTGCAAT ATATCTTGCA 2800
 AGCCGCGCCC GAAGATCGCC GCACGCTGAT GTTCTCGGCC ACCGTGCCGC 2850
 GCGAGATTGA AAAACTGGCC CGCGACTTCC AAAATGACGC CCTGCGTCTG 2900
 40 GAAACCCGTG GCGAGGCCAA GCAGACAAC GACATCAGCT ACCAAGCTTT 2950
 GTCGGTCACC ATGCGCGATC GCGAAAACGC CATTTTCAAC ATGCTGCGTT 3000

- 3 -

TTTATGAATC GCGCACGGCG ATCATCTTCT GCAAGACCCG CGCCAATGTG 3050
 AATGATCTGC TGTCGCGGAT GAGCGGTCGT GGCTTCCGCG TGGTGGCCCT 3100
 GTCGGGCGAG CTGTCGCAAC AGGAACGCAC CAACGCGCTG CAAGCGCTGC 3150
 GTGATGGCCG CGCCAACGTT TGTATCGCGA CCGACGTCGC GGC GCGCGGC 3200
 5 ATTGACTTGC CGGGCCTCGA GCTGGTGATC CACTACGATC TGCCGACCAA 3250
 TGCCGAAACC CTGCTGCACC GCTCGGGCCG TACCGGCCGC CGGGTGCCAA 3300
 GGGCGTCTCG GCGCTGATCG TCACCCCCGG CGATTTCAAA AAAGCGCAGC 3350
 GTTTGCTGAG CTTTGCCAAA GTGACCGCGG AATGGGGCAA GGCGCCTTCG 3400
 GCCGAAGA 3408

10

<210> 2

<211> 609

<212> PRT

15 <213> Gluconobacter oxydans

<220>

<221> SIGNAL

<222> (1) .. (31)

<220>

20 <221> CHAIN

<222> (32) .. (609)

<400> 2

Met Leu Pro Lys Ser Leu Lys His Lys Asn Gly Ala Met Arg Leu 15
 25 Val Ala Ala Ser Thr Leu Ala Leu Met Ile Gly Ala Gly Ala His 30
 Ala Gln Val Asn Pro Val Glu Val Pro Val Gly Ala Asn Glu Thr 45
 30 Phe Thr Ser Arg Val Leu Thr Thr Gly Leu Ser Asn Pro Trp Glu 60
 Ile Thr Trp Gly Pro Asp Asn Met Leu Trp Val Thr Glu Arg Ser 75
 Ser Gly Glu Val Thr Arg Val Asp Pro Asn Thr Gly Glu Gln Gln 90
 35 Val Leu Leu Thr Leu Thr Asp Phe Ser Val Asp Val Gln His Gln 105
 Gly Leu Leu Gly Leu Ala Leu His Pro Glu Phe Met Gln Glu Ser 120
 40 Gly Asn Asp Tyr Val Tyr Ile Val Tyr Thr Tyr Asn Thr Gly Thr 135
 Glu Glu Ala Pro Asp Pro His Gln Lys Leu Val Arg Tyr Ala Tyr 150

- 4 -

Asp Ala Ala Ala Gln Gln Leu Val Asp Pro Val Asp Leu Val Ala 165
Gly Ile Pro Ala Gly Asn Asp His Asn Gly Gly Arg Ile Lys Phe 180
5 Ala Pro Asp Gly Gln His Ile Phe Tyr Thr Leu Gly Glu Gln Gly 195
Ala Asn Phe Gly Gly Asn Phe Arg Arg Pro Asn His Ala Gln Leu 210
10 Leu Pro Thr Gln Glu Gln Val Asp Ala Gly Asp Trp Val Ala Tyr 225
Ser Gly Lys Ile Leu Arg Val Asn Leu Asp Gly Thr Ile Pro Glu 240
Asp Asn Pro Glu Ile Glu Gly Val Arg Ser His Ile Phe Thr Tyr 255
15 Gly His Arg Asn Pro Gln Gly Ile Thr Phe Gly Pro Asp Gly Thr 270
Ile Tyr Ala Thr Glu His Gly Pro Asp Thr Asp Asp Glu Leu Asn 285
20 Ile Ile Ala Gly Gly Gly Asn Tyr Gly Trp Pro Asn Val Ala Gly 300
Tyr Arg Asp Gly Lys Ser Tyr Val Tyr Ala Asp Trp Ser Gln Ala 315
Pro Ala Asp Gln Arg Tyr Thr Gly Arg Ala Gly Ile Pro Asp Thr 330
25 Val Pro Gln Phe Pro Glu Leu Glu Phe Ala Pro Glu Met Val Asp 345
Pro Leu Thr Thr Tyr Trp Thr Val Asp Asn Asp Tyr Asp Phe Thr 360
30 Ala Asn Cys Gly Trp Ile Cys Asn Pro Thr Ile Ala Pro Ser Ser 375
Ala Tyr Tyr Tyr Ala Ala Gly Glu Ser Gly Ile Ala Ala Trp Asp 390
Asn Ser Ile Leu Ile Pro Thr Leu Lys His Gly Gly Ile Tyr Val 405
35 Gln His Leu Ser Asp Asp Gly Gln Ser Val Asp Gly Leu Pro Glu 420
Leu Trp Phe Ser Thr Gln Asn Arg Tyr Arg Asp Ile Glu Ile Ser 435
40 Pro Asp Asn His Val Phe Val Ala Thr Asp Asn Phe Gly Thr Ser 450
Ala Gln Lys Tyr Gly Glu Thr Gly Phe Thr Asn Val Leu His Asn 465

- 5 -

Pro Gly Ala Ile Leu Val Phe Ser Tyr Val Gly Glu Asp Ala Ala 480
Gly Gln Thr Gly Met Met Thr Ala Pro Ala Pro Gln Thr Gln Tyr 495
5 Thr Gln Val Pro Ala Glu Gly Ala Gly Ala Gly Ala Thr Glu Val 510
Ala Asp Val Asp Tyr Asp Thr Leu Phe Thr Glu Gly Gln Thr Leu 525
10 Tyr Gly Ser Ala Cys Ala Ala Cys His Gly Ala Ala Gly Gln Gly 540
Ala Gln Gly Pro Thr Phe Val Gly Val Pro Asp Val Thr Gly Asp 555
Lys Asp Tyr Leu Ala Arg Thr Ile Ile His Gly Phe Gly Tyr Met 570
15 Pro Ser Phe Ala Thr Arg Leu Asp Asp Glu Glu Val Ala Ala Ile 585
Ala Thr Phe Ile Arg Asn Ser Trp Gly Asn Asp Glu Gly Ile Leu 600
20 Thr Pro Ala Glu Ala Ala Ala Thr Arg 609

<210> 3
<211> 14
25 <212> PRT
<213> Gluconobacter oxydans

<400> 3
Gln [Xaa/Gly] Asn [Pro/Lys] Val Glu Val Pro Val Gly Ala Asn Glu Thr 14
30

<210> 4
<211> 31
<212> PRT
35 <213> Gluconobacter oxydans

<220>
<221> SIGNAL
<222> (1) .. (31)
40
<400> 4

- 6 -

Met Leu Pro Lys Ser Leu Lys His Lys Asn Gly Ala Met Arg Leu 15
Val Ala Ala Ser Thr Leu Ala Leu Met Ile Gly Ala Gly Ala His 30
5 Ala 31

<210> 5
<211> 578
10 <212> PRT
<213> Gluconobacter oxydans

<220>
<221> CHAIN
15 <222> (1)..(578)

<400> 5
Gln Val Asn Pro Val Glu Val Pro Val Gly Ala Asn Glu Thr Phe 15
20 Thr Ser Arg Val Leu Thr Thr Gly Leu Ser Asn Pro Trp Glu Ile 30
Thr Trp Gly Pro Asp Asn Met Leu Trp Val Thr Glu Arg Ser Ser 45
Gly Glu Val Thr Arg Val Asp Pro Asn Thr Gly Glu Gln Gln Val 60
25 Leu Leu Thr Leu Thr Asp Phe Ser Val Asp Val Gln His Gln Gly 75
Leu Leu Gly Leu Ala Leu His Pro Glu Phe Met Gln Glu Ser Gly 90
30 Asn Asp Tyr Val Tyr Ile Val Tyr Thr Tyr Asn Thr Gly Thr Glu 105
Glu Ala Pro Asp Pro His Gln Lys Leu Val Arg Tyr Ala Tyr Asp 120
Ala Ala Ala Gln Gln Leu Val Asp Pro Val Asp Leu Val Ala Gly 135
35 Ile Pro Ala Gly Asn Asp His Asn Gly Gly Arg Ile Lys Phe Ala 150
Pro Asp Gly Gln His Ile Phe Tyr Thr Leu Gly Glu Gln Gly Ala 165
40 Asn Phe Gly Gly Asn Phe Arg Arg Pro Asn His Ala Gln Leu Leu 180

- 7 -

Pro Thr Gln Glu Gln Val Asp Ala Gly Asp Trp Val Ala Tyr Ser 195

Gly Lys Ile Leu Arg Val Asn Leu Asp Gly Thr Ile Pro Glu Asp 210

5 Asn Pro Glu Ile Glu Gly Val Arg Ser His Ile Phe Thr Tyr Gly 225

His Arg Asn Pro Gln Gly Ile Thr Phe Gly Pro Asp Gly Thr Ile 240

Tyr Ala Thr Glu His Gly Pro Asp Thr Asp Asp Glu Leu Asn Ile 255

10 Ile Ala Gly Gly Gly Asn Tyr Gly Trp Pro Asn Val Ala Gly Tyr 270

Arg Asp Gly Lys Ser Tyr Val Tyr Ala Asp Trp Ser Gln Ala Pro 285

15 Ala Asp Gln Arg Tyr Thr Gly Arg Ala Gly Ile Pro Asp Thr Val 300

Pro Gln Phe Pro Glu Leu Glu Phe Ala Pro Glu Met Val Asp Pro 315

Leu Thr Thr Tyr Trp Thr Val Asp Asn Asp Tyr Asp Phe Thr Ala 330

20 Asn Cys Gly Trp Ile Cys Asn Pro Thr Ile Ala Pro Ser Ser Ala 345

Tyr Tyr Tyr Ala Ala Gly Glu Ser Gly Ile Ala Ala Trp Asp Asn 360

25 Ser Ile Leu Ile Pro Thr Leu Lys His Gly Gly Ile Tyr Val Gln 375

His Leu Ser Asp Asp Gly Gln Ser Val Asp Gly Leu Pro Glu Leu 390

Trp Phe Ser Thr Gln Asn Arg Tyr Arg Asp Ile Glu Ile Ser Pro 405

30 Asp Asn His Val Phe Val Phe Thr Asp Asn Ile Thr Thr Thr Thr 420

Gln Lys Tyr Gly Glu Thr Gly Phe Thr Asn Val Leu His Asn Pro 435

35 Gly Ala Ile Leu Val Phe Ser Tyr Val Gly Glu Asp Ala Ala Gly 450

Gln Thr Gly Met Met Thr Ala Pro Ala Pro Gln Thr Gln Tyr Thr 465

Gln Val Pro Ala Glu Gly Ala Gly Ala Gly Ala Thr Glu Val Ala 480

40 Asp Val Asp Tyr Asp Thr Leu Phe Thr Glu Gly Gln Thr Leu Tyr 495

- 8 -

Gly Ser Ala Cys Ala Ala Cys His Gly Ala Ala Gly Gln Gly Ala 510
Gln Gly Pro Thr Phe Val Gly Val Pro Asp Val Thr Gly Asp Lys 525
5 Asp Tyr Leu Ala Arg Thr Ile Ile His Gly Phe Gly Tyr Met Pro 540
Ser Phe Ala Thr Arg Leu Asp Asp Glu Glu Val Ala Ala Ile Ala 555
10 Thr Phe Ile Arg Asn Ser Trp Gly Asn Asp Glu Gly Ile Leu Thr 570
Pro Ala Glu Ala Ala Ala Thr Arg 578

15 <210> 6
<211> 17
<212> DNA
<213> Artificial Sequence
20 <220>
<223> an artificially synthesized primer sequence

<400> 6
carggyaacc csgtbga 17

25 <210> 7
<211> 17
<212> DNA
<213> Artificial Sequence

30 <220>
<223> an artificially synthesized primer sequence

35 <221> misc_feature
<222> 9
<223> n is a or g or c or t

<400> 7
40 gtytcgttng crccvac 17

- 9 -

<210> 8
<211> 15
<212> DNA
5 <213> Artificial Sequence

<220>
<223> an artificially synthesized primer sequence

10 <400> 8
cagggtacc cggtc 15

<210> 9
15 <211> 15
<212> DNA
<213> Artificial Sequence

<220>
20 <223> an artificially synthesized primer sequence

<400> 9
gactcgtttg cgccc 15